Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal 2021, 15(3), 167–178
https://doi.org/10.12913/22998624/140542
ISSN 2299-8624, License CC-BY 4.0

Received: 2021.06.14
Accepted: 2021.07.20
Published: 2021.08.16

Classification of Seizure Types Using Random Forest Classifier

Ashjan Basri[1*], Muhammad Arif[1]

[1] Department of Computer Science, College of Computer and Information Systems, Umm Alqura University, Saudi Arabia

* Corresponding author's e-mail: syedmarif2003@yahoo.com

ABSTRACT

Epilepsy is one of the most common mental disorders in the world, affecting 65 million people. The prevalence in Arab countries of Epilepsy is estimated at 174 per 100,000 individuals, and in Saudi Arabia is 6.54 per 1,000 individuals. Epilepsy seizures have different types, and each patient needs to have a treatment plan according to the seizure type. Hence, accurate classification of seizure type is an essential part of diagnosing and treating epileptic patients. In this paper, the features based on fast Fourier transform from EEG montages were used to classify different types of seizures. Since the distribution of classes was not uniform, the dataset suffered from severe imbalance. Various algorithms were used to under-sample the majority class and over-sample the minority classes. Random forest classifier produced classification accuracy of 96% to differentiate three types of seizures from the healthy EEG reading.

Keywords: EEG, fast fourier transform, seizure, random forest.

INTRODUCTION

Epilepsy is a pathological condition that appears because of abnormal electrical activity in the brain [1]. It is one of the significant issues affecting approximately 65 million people worldwide, comprising 1% of the world population [2]. The incidence of epilepsy is estimated at 174 per 100,000 individuals in Arab countries. The prevalence of epilepsy in the Kingdom of Saudi Arabia is 6.54 per 1,000 individuals [3]. However, medical care services are not available to one-third of epileptic patients. They must find ways to live and manage their everyday lives. The quality of medical treatment is not up to the mark even if the medical care services are available to epileptic patients [4]. Diagnosis and treatment of epileptic patients depend on the type of seizures [4]. Electroencephalography (EEG) recording is one of the techniques neurologists use to analyze the abnormalities of brainwave functions. It has been widely used for many years to diagnose the brain conditions such as sleep disorders, dementia, epilepsy, etc., [5]. A standard EEG examination lasts approximately 20 to 30 minutes. This period is not always sufficient to record seizures for patients. Fifty percent of the patients need to conduct several EEG sessions in their initial diagnosis. It may take many hours to several days, making the EEG test incredibly costly [6]. The EEG recording is analyzed manually by a qualified physician to finalize the diagnosis.

Additionally, the perception of the EEG primarily depends on the reader's subjective evaluation, which may miss events or leads to misdiagnosis [6]. In this context, many existing studies in the literature applied the Machine Learning techniques to classify the abnormality in the EEG data [7, 8, 9]. A clinical decision support system based on Machine Learning algorithms to perform the EEG interpretation task can assist the neurologists in accelerating the diagnosis process. Epilepsy is a symptom of many neurological disorders. Several studies define epilepsy as a brain disorder with frequent seizures, which appears in various forms and symptoms. It could be in the form of repeated or constant motor activity in a specific area or all parts of the body [10, 1, 4, 2]. Epilepsy is generally diagnosed after at least two seizures not caused by an existing physical issue, such as low blood sugar or medicine [11]. Seizures are a sudden disturbance of electrical activity in the brain that causes abnormal and unbalanced movements without the patient's consciousness [5]. These movements may occur as muscle spasms, muscle contractions, muscle relaxations, or frequent limb vibrations.

In epilepsy, a seizure may also appear in the changes in the sensations or enhance autonomy nervous system function. If seizures occur from a particular region of the brain, they may be visible in the motor function controlled by that brain region [3]. For example, in the brain area that controls the movement of the thumb from the right side of the brain, a seizure can start when the left thumb or hand jerks. There are medical conditions having the symptoms similar to epilepsy that make the diagnosis difficult. Patients may be required to do some tests, including Magnetic Resonance Imaging (MRI), Computerized Tomography (CT), or Electroencephalogram (EEG) test to ensure a correct diagnosis. EEG provides the potential cause of epilepsy and the seizure type of a patient [6]. The main focus of this study was to analyze the EEG data and predict the seizure types.

Seizures are classified into two broad categories based on the patient's behavior during the seizure and the brain activity [11]. Generalized seizures affect the entire brain and are also known as Grand-Mal convulsions. The patient usually loses his/her conscience and collapses during this kind of seizure. It starts with a general body stiffening known as the Tonic phase, which generally lasts for 30 or 60 seconds. Then, the patient goes through violent jerking, known as the Clonic phase, which also lasts for 30 to 60 seconds. Afterwards, the patient goes to deep sleep, which is a post-seizure phase. Injuries such as tongue biting and urinary incontinence can occur during Grand-Mal seizure.

Generalized seizures have six types: Generalized Tonic-Clonic, Absence, Myoclonic, Clonic, Tonic, and Atonic. Each one of these types has different symptoms [12]. Focal or Partial seizures affect only one part of the brain, but sometimes spread into other brain lobes. Partial seizures have two types: Simple and Complex type. In a simple partial seizure, patients preserve their consciousness, while during a complex partial seizure, patients lose their consciousness [13, 11]. This diversity of symptoms occurs due to the differences in the source of electrical activities in the brain.

Moreover, this diversity leads to the declassification of the epilepsy type and misdiagnosis. An electroencephalogram (EEG) test identifies

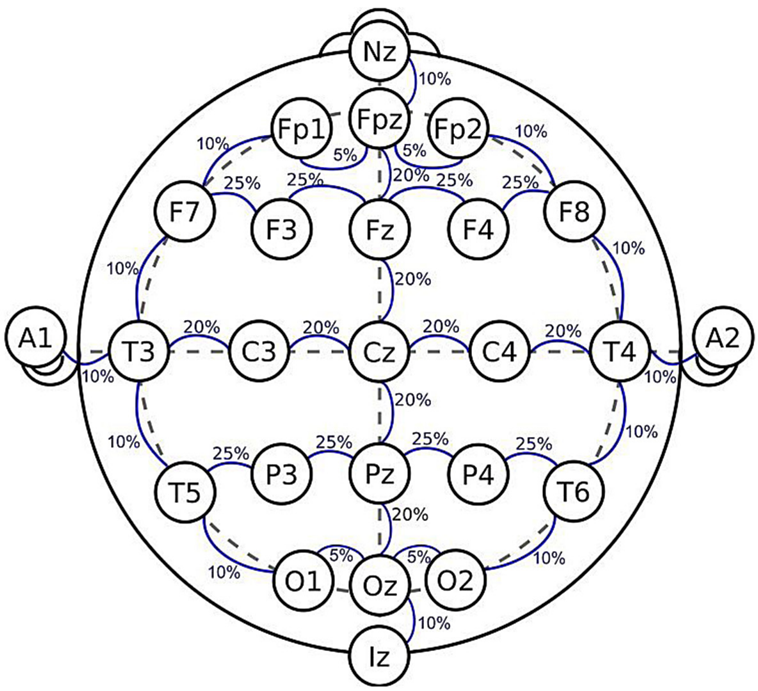


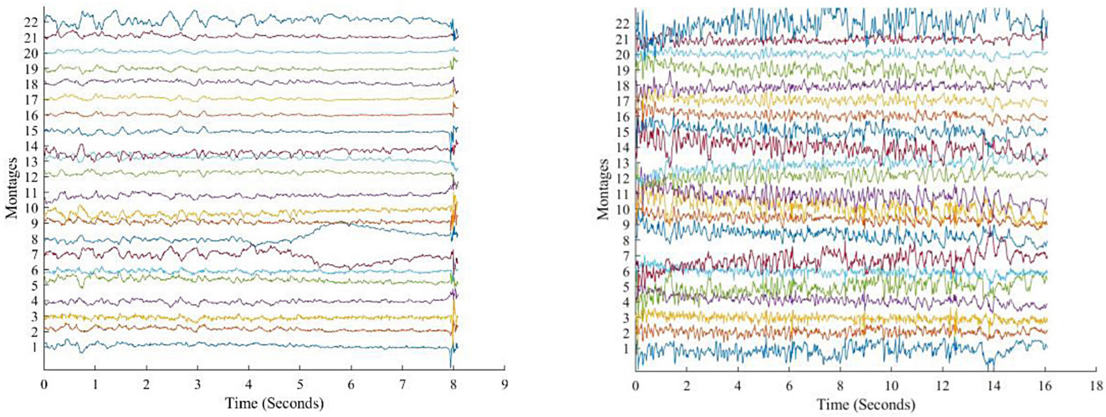
Fig. 1. Standard 10–20 system for electrode placement [14]

abnormal brain activities and tracks patterns of the brain waves. Electrodes are placed on the head of the patient, and electrical signals are sensed through these electrodes. A physician decides whether the brain's electrical activity is normal or abnormal by looking at the electrical activity.

Typically, the placement of electrodes follows the International 10/20 system, which is 21 electrodes on the scalp are evenly distributed as seen in Figure 1, with the distance between electrodes is either 10% or 20% of the total distance from nasion (front) to inion (back) [14]. Each electrode begins with a letter corresponding to the region where the electrode is placed.

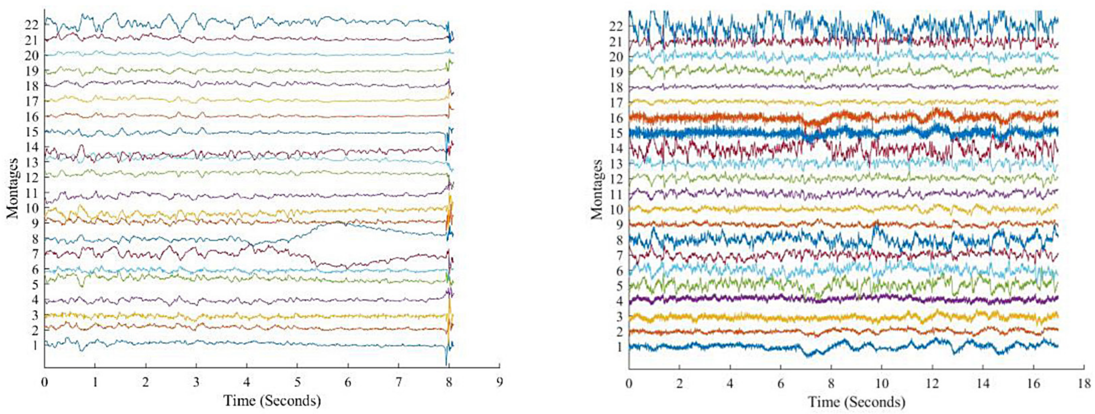
The frequency analysis of EEG signal shows five frequency bands: delta, theta, alpha, beta, and gamma [15, 16].

These frequency rhythms reflect the state of brain activity. Normal EEG of adults is consistent and low amplification activity in the alpha or beta ranges. EEG also shows artifacts associate with eye blinks/movements and muscle artifacts [17]. Several distinct features should be present in EEG to characterize a seizure event: evolution, spike, or polyspic morphology of the slow-wave, rhythm, synchrony, continuity, and frequency. A seizure should have a duration of at least 10 seconds and no longer than three seconds [17]. If there is a gap of more than 3 seconds, events are taken as two separate seizures. The most well-known type of epilepsy is a generalized seizure, which covers a greater number of channels and areas of the patient's skull, as shown in Figure 2. Figures 2(a) and 3(a) show normal EEG activity in all montages. Figure 2(b) shows the montages obtained from the EEG channels of a person suffering from generalized seizure (GNSZ). It can be observed that abnormal EEG patterns are visible in most of the channels. In the case of non-specific focal seizures, it covers a broad range of seizure etiologies. As such, it has a significant variation in appearance, length, and focal. The



(a) Normal EEG (b) EEG of GNSZ

Fig. 2. Montages from EEG channels for a Normal person and person suffering from generalized seizure (GNSZ)



(a) Normal EEG (b) EEG of FNSZ

Fig. 3. Montages from EEG channels for a Normal person and person suffering from localized seizure (FNSZ)

primary indicator of a non-specific seizure event is morphology (as shown in Figure 3). Figure 3(b) shows the montages of a person suffering from FNSZ. In the case of localized seizures, having a focal point reflects as abnormal EEG patterns in few montages.

Related Work

This section summarizes the use of the Machine Learning algorithms in the analysis of the EEG signals in epilepsy disorders. The main focus are the epilepsy feature selection methods, epilepsy seizure detection, and epilepsy prediction strategies. Niknazar et al. [18] proposed the feature extraction strategies for automatic epileptic EEG wave recognition, and various machine learning models were used to differentiate epileptic seizure and non-seizure events. Decision Tree (DT), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Nearest Neighbour (NN) classifiers are used in the study. One-dimensional Local Gradient Pattern (1D-LGP) and Local Neighbour Descriptive Pattern (LNDP) feature extraction strategies with ANN Classification algorithm achieved an average classification accuracy of 99.80%. In another paper, Niknazar et al. [19] proposed early detection of seizure using the EEG data. Features are extracted from intracranial EEG waves obtained through intrusive pre-surgical epilepsy screening of the patients with drug-resistant focal seizures. Time and space domain features are obtained, and correlation analysis quantifies the characteristics of the EEG indicators. Ictal states are distinguished from pre-ictal conditions successfully(p ¡ 0.01) [19]. Acharya et. [20] used a convolutional neural network comprising 13 layers and attained an accuracy of 90% on five subjects. Yang et al. [21] proposed a novel feature extraction technique called MinMaxHist to measure the topological patterns of EEG. They have used MinMaxHist features and other time-domain features to classify the epileptic event. The classification accuracy of 86% is achieved from 30 features while testing on patient-independent studies. The diversity of epileptic seizures makes it challenging to distinguish the sequence of epilepsy episodes from normal EEG. Wulandari et al. [22] considered the EEG signal spectrum characteristics to differentiate seizure and non-seizure disorders of epilepsy. A technology that could predict the onset of the seizure event can alert both the patient and physician, and the quality of life of the patient can be improved significantly. Deriche et al. [23] suggested the features using Singular Value Decomposition (SVD) of the EEG signal Time-Frequency matrix. These features can distinguish between regular and abnormal EEG traces (both ictal and interictal). They claimed the classification accuracy of more than 99% in binary classification on ten subjects only. Dash et al. [24] proposed an iterative filtering-based decomposition of EEG signals to improve the accuracy of seizure detection. Hidden Markov Model is used as a probabilistic classifier to detect the seizure event. The feature set includes power spectral density, time-domain features, dynamic mode decomposition power, variance, and Katz fractal dimension. The classification accuracy was about 99% on two datasets containing 40 subjects in total. Sharma et al. [25] used localized wavelet filter banks to detect the abnormal EEG signals. They have found a classification accuracy of 78% using an SVM classifier. Ihsan Ullah et al. [26] trained an ensemble of one-dimensional convolutional neural networks to classify the normal and abnormal EEG patterns by majority voting. Classification accuracy based on 10-folds cross-validation for two classes was 96% and 99% for three classes. Nkengfack et al. [27] proposed EEG rhythms decomposition-based Jacobi polynomial transforms (JPTs) to extract various time and frequency domain features. Relevant features were computed using linear discriminant analysis. Support vector machine classifier produced classification accuracy in the range of 97% to 100% for various experiments. Alickovic et al. [28] tried different combinations of pre-processing techniques, feature extraction, and classifiers on two EEG datasets. They have found that the Random Forest classifier performed best on the statistical feature set extracted by applying wavelet packet decomposition. Classification accuracy for three-class problems was more than 99% for both datasets. Few works demonstrated the methodologies to predict the epileptic seizure onset ahead of time. Wei et al. [29] proposed a long-term recurrent convolutional network to predict the seizures. Convolution layers extract the deep features from the data, and long short- term memory (LSTM) layers learn the time sequences in the data. The model predicted the seizures with 93% accuracy. Usman et al. [30] pre-processed the EEG signal by empirical mode decomposition to remove the noise. Generative Adversarial Networks (GAN) generated more samples to deal with the class imbalance problem. Features are extracted from

CNN, and classification is done through LSTM units. The proposed method achieved an accuracy of 93% with an average onset time of 32 minutes. Khan et al. [31] pre-processed the EEG signal using wavelet transform. After the pre-processing step, convolutional filters are used to learn the interictal, preictal, and ictal events. The preictal phase occurs about ten minutes before the start of the seizure. The CNN classifier identified the preictal stage with 87% accuracy. A similar work by Truong et al. [32] used a deep learning model (CNN) on intracranial and scalp electroencephalogram (EEG) datasets. They have achieved a sensitivity of 81% in predicting the preictal and interictal segments on these datasets.

MATERIALS AND METHODS

The EEG dataset of The Temple University Hospital (TUH) [34, 35] was used in this paper. The corpus is an open-source database [33] and contains scalp EEG signal recordings. The sampling rate varies in the EEG recordings. A variation of 40 different channel configurations with sample frequency ranges from 250 Hz to 1024 Hz. The database contains two folders: Dev and Train. The data present in the Train folder were used. It covers eight types of seizures, as given in Table 1. The dataset includes a total of 2377 seizures from more than 200 patients. The folder (01_tcp_ar) from the train dataset folder was picked, which is the average reference configuration and annotations use the TCP channel configuration. The details about the corpus can be found in [34, 35].

Pre-processing of EEG data

The sampling frequency is different for different sessions. More than 90% of sessions are recorded with a sampling frequency of 250 Hz. All signals sampled on a sampling frequency different from 250 Hz are resampled to 250 Hz. Resampling of the signal is done through a Least-squares linear-phase FIR filter design [36, 37]. Although a different set of channels are used in the sessions, all sessions include the International 10/20 system. After resampling all the sessions to one sampling frequency, the EEG signals are converted into a set of montages or differentials as described in Table 2. A total of 22 montages are extracted from the EEG channels measures in the 10/20 system.

Table 1. The details of Seizure Type in TUH database v1.5.2

Seizure Type	Number of Seizures	Percentage
Focal Non-Specific (FNSZ)	1536	65%
Generalized Non-Specific (GNSZ)	409	17%
Complex Partial (CPSZ)	283	12%
Absence (ABSZ)	50	2%
Tonic (TNSZ)	18	1%
Tonic Clonic (TCSZ)	30	1%
Simple Partial (SPSZ)	49	2%
Myoclonic (MYSZ)	2	0%

Table 2. Montages definition

Montage	Differential of Channels	Montage	Differential of Channels
0	FP1-F7	11	CZ-C4
1	F7-T3	12	C4-T4
2	T3-T5	13	T4-A2
3	T5-O1	14	FP1-F3
4	FP2-F8	15	F3-C3
5	F8-T4	16	C3-P3
6	T4-T6	17	P3-O1
7	T6-O2	18	FP2-F4
8	A1-T3	19	F4-C4
9	T3-C3	20	C4-P4
10	C3-CZ	21	P4-O2

Feature Extraction

Discrete Fourier Transform (DFT) transforms a time series from time-domain to frequency-domain. The Fast Fourier Transform (FFT) is a computationally efficient way of computing DFT.

Let $x_0, x_1, \ldots, x_{N-1}$ is a time series. The Discrete Fourier Transform can be calculated by equation (1) as follows:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi k/N} \quad k = 0,1,\ldots,N-1 \qquad (1)$$

N-points FFT were applied on each of the montage signals and concatenated as a feature vector.

Methods to treat Data Imbalance

In many real-world applications, the normal class dominates the abnormal courses, and hence a bias in the classifier training produces poor classification performance. There are numerous publications handling the class imbalance problem. There are two ways to tackle this problem. Either the majority class is under-sampled to reduce the number of instances in the majority class, oversample the minority class to increase its instances, or both. The under-sampling methods may include a random selection of instances, selection of instances based on some criteria [38, 39, 40], selection of borderline instances [41], considering the overlapping regions of minority and majority class belonging to the minority class [42], etc. In turn, the oversampling techniques for minority classes include Synthetic Minority Over-sampling Technique (SMOTE) [43], borderline SMOTE [44], Adaptive synthetic sampling approach for imbalanced learning (ADASYN) [45], etc. In SMOTE, a minority class is selected randomly and finds K nearest neighbours of this instance. One neighbour is chosen randomly from these nearest neighbours, and a new instance is generated between the instance and the selected nearest neighbour. In turn, in the ADASYN algorithm, the number of synthetic instances generated depends on the instance that is difficult to learn. In borderline SMOTE, only those instances from the minority classes are over-sampled near the classification boundary. A good survey can be found in [46].

Random Forest Classifier

Random Forest (RF) classifier [47] is a supervised classifier based on an ensemble of decision trees. The ensemble of trees is training by the bagging method. Merging the classification of the ensemble of the decision trees improves the overall classification performance. Different hyper-parameters associated with the RF classifier can be optimized to maximize the classification accuracy. These parameters are the number of decision trees, the maximum number of features to split the node, and the minimum number of leaves to split the node.

Performance Measures

The classifier performance can be expressed in terms of accuracy, precision, Recall, and F-measure. Precision is defined as the number of truly positive instances (TP: True Positive) divided by the total number of positive instances (including True Positive and False Positive).

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

Recall or Sensitivity is defined as the number of instances truly classified as positive (TP: True Positive) divided by the total Positive instances including TP and FN (False Negative).

$$Recall = \frac{TP}{TP + FN} \qquad (3)$$

F-measure is the combination of precision and Recall and defined as:

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \qquad (4)$$

RESULTS AND DISCUSSIONS

A window size of 10 seconds contains 2500 data points. A 128-points FFT is applied to mutually exclusive windows of 10 seconds in all 22 montages, giving a feature set of 1430 points. Table 3 shows the class distribution of all eight classes. Since the number of TNSZ, TCSZ, and SPSZ is very low, these three classes were ignored and only four classes were retained for classification. It is evident from table 3 that the Normal class dominates all other classes creating severe class imbalance. Therefore, it is essential to balance the class distribution before applying any classifier. First, the majority class was treated by under-sampling this class and only important instances were kept. The under-sampling algorithm is described in Table 4. This algorithm is similar to the concept of support vectors in the support vector machines. In the support vector machine, support vectors are calculated that are most important in the sense of decision boundary [70]. The algorithm finds the nearest neighbours of the majority class from the instances of the minority classes that lie near the decision boundary. These instances are near the decision boundary and play a crucial role in deciding the classification boundary by the classifier.

Figure 4 shows the class distribution of four classes after applying the under-sampling algorithm. Although the majority class instances are reduced, it still dominates other classes (Normal 75%, GNSZ 7.9%, FNSZ 14%, CNSZ 3.1%). Random Forest classifier is applied to the dataset, and performance of the classier for 10-folds cross-validation is listed in Table 5 as Precision, Recall, F-measure, and area under the ROC (AUC). The normal class has shown higher values for precision (0.84) and Recall (0.992) but a low AUC value. The same is true for other classes. The Recall values for GNSZ, FNSZ, and CNSZ are meager, showing the poor performance of the classifier in recognizing these classes. Table 6 shows the confusion matrix for all four classes. It is evident from the table that

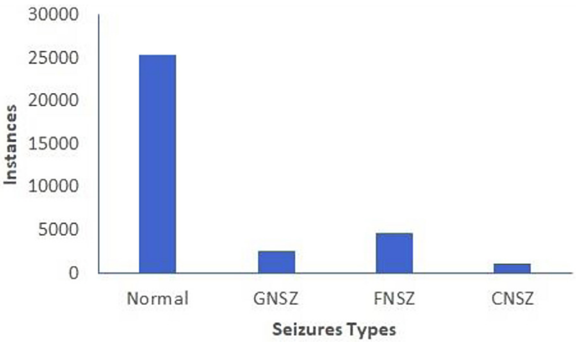


Fig. 4. Distribution of the Classes after under-sampling the majority class

Table 3. Distribution of Classes

Class	Instances	Percentage
Normal	161886	94.95
GNSZ	2596	1.52
FNSZ	4696	2.75
CNSZ	1051	0.62
TNSZ	173	0.10
TCSZ	52	0.03
SPSZ	36	0.02

Table 4. Under-sampling Algorithm for Majority Class

Algorithm	Under-sampling Algorithm
Inputs	
M	Majority Class
C_i	Minority Classes
K	Number of Nearest Neighbors
Ø	List of important Majority Class instances
KNN	Function to find K-nearest neighbors
Step 1	for i=1 to number of minority classes
Step 2	For j=1 to instances in class C_i
Step 3	For each instance j in Class C_i find *K* nearest instances in Majority Class *M* using function *KNN*
Step 4	Put *K* instances in *Ø*
Step 5	end j
Step 6	end i
Step 7	Remove redundant instances from *Ø*
Outputs:	Under-sampled Majority Class *Ø*

Table 5. Performance analysis of random forest classifier on an under-sampled dataset

Class	Precision	Recall	F-Measure	AUC
Normal	0.814	0.992	0.894	0.474
GNSZ	0.891	0.446	0.595	0.612
FNSZ	0.938	0.204	0.335	0.407
CNSZ	0.978	0.430	0.597	0.642
Weighted Average	0.842	0.823	0.784	0.480

Table 6. Confusion matrix of random forest classifier on an under-sampled dataset

Class	Normal	GNSZ	FNSZ	CNSZ
Normal	25132	139	47	10
GNSZ	1423	1158	15	0
FNSZ	3738	2	956	0
CNSZ	598	0	1	452

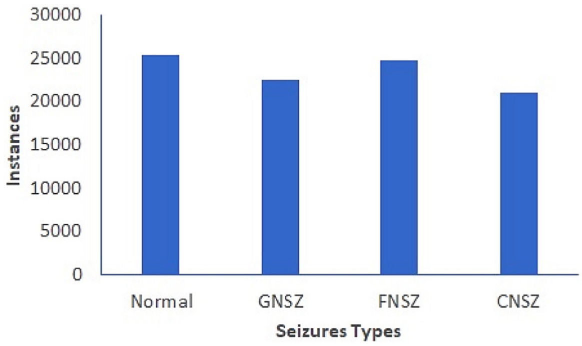


Fig. 5. Distribution after Over-sampling Minority Classes using SMOTE

many instances of minority classes are confused with the majority class of Normal. Due to this reason, the Recall values of these classes are low. The overall classification accuracy is 82%.

In order to reduce the class imbalance further, oversampling algorithms are applied to the minority classes. Over-sampling algorithms are used to create additional instances of minority classes. In this paper, two algorithms were tried, SMOTE and ADAYSN. The new 16000 instances are created for three minority classes (GNSZ, FNSZ, CNSZ) using SMOTE algorithm. The distribution of all four classes after oversampling through SMOTE is plotted in Figure 5. The number of instances in each class shows that now all the classes have almost equal number of instances.

Random Forest is applied to this over-sampled dataset, and the performance of the classifier is obtained by ten-fold cross-validation. Table 7 explains the performance of the classifier of the new dataset.

Precision and Recall for all the classes are above 90%, which shows better classification than the dataset before oversampling. Overall classification accuracy is 96%. The RF classifier showed excellent performance on minority classes as well. A weighted average of Precision and Recall is 0.96, and AUC is 0.996. The confusion matrix in Table 8 shows that there are still few instances that are confused with Normal Class.

Similarly, few instances of the Normal class are also confused with other classes. The CNSZ class has the least number of instances confused with other classes, whereas many instances of the Normal class are confused with the FNSZ class. Oversampling the minority classes improved the ability of the classifier to find the best decision boundary among all the classes. Hence, the random classifier trained on an over-sampled training dataset helps optimize the classifier parameters.

The adaptive synthetic (ADASYN) over-sampling method is different from SMOTE as ADASYN uses the density distribution information to generate new instances. It can also compensate for the skewed density distribution of the classes. Hence, ADASYN is used to over-sample the minority classes, and the performance of the random forest classifier is compared to the performance of the random forest classifier trained on an SMOTE-based over-sampled dataset.

Random Forest classifier is applied to the over-sampled dataset, and performance measures are tabulated in the Table 9 for ten-folds cross-validation. Precision, Recall, F-measure, and AUC for all the classes are similar to the performance of the RF classifier on the over-sampled dataset by SMOTE. The overall classification accuracy is 96.2%. The confusion matrix for all four

Table 7. Performance analysis of random forest classifier (Over-sampled by SMOTE)

Class	Precision	Recall	F-Measure	AUC
Normal	0.945	0.910	0.927	0.988
GNSZ	0.969	0.975	0.972	0.998
FNSZ	0.937	0.966	0.951	0.997
CNSZ	0.995	0.998	0.997	1.000
Weighted Average	0.960	0.960	0.960	0.996

Table 8. Confusion matrix of Random Forest classifier (Over-sampled by Smote)

Class	Normal	GNSZ	FNSZ	CNSZ
Normal	23045	673	1518	92
GNSZ	468	22115	109	0
FNSZ	831	34	24523	4
CNSZ	35	0	9	21058

Table 9. Performance analysis of Random Forest classifier (Over-sampled by ADASYN)

Class	Precision	Recall	F-Measure	AUC
Normal	0.961	0.898	0.928	0.989
GNSZ	0.964	0.986	0.975	0.998
FNSZ	0.933	0.972	0.952	0.997
CNSZ	0.995	1.000	0.997	1.000
Weighted Average	0.962	0.962	0.961	0.996

Table 10. Confusion matrix of random forest classifier (Over-sampled by ADASYN)

Class	Normal	GNSZ	FNSZ	CNSZ
Normal	22740	796	1683	109
GNSZ	261	22343	64	1
FNSZ	655	38	24384	2
CNSZ	2	0	3	21060

classes is illustrated in Table 10. It is evident from the table that more instances of Normal class and FNSZ class are confused with each other. Tables 7 and 9 presents classification results when two different types of over-sampling methods (SMOTE and ADASYN) are used. Both over-sampling methods have shown similar results, and ADASYN based over-sampling is slightly better than SMOTE.

CONCLUSIONS

This paper focused on classifying various types of seizures from the healthy brain electrical activity using EEG data. It is essential to diagnose abnormal brain activity and the type of seizures affecting brain activity. Three types of epileptic seizures were considered and successfully classified against regular brain electrical activity. As of many real-life scenarios, the number of epilepsy patients is few as compared to healthy subjects. Therefore, the healthy class (a person with normal EEG) dominates the seizure classes in terms of the number of instances. The majority class was under-sampled by an under-sampling algorithm, and minority classes were over-sampled by two algorithms before applying the Random Forest Classifier. It is evident from the results that the Random Forest classifier was able to classify all four classes with an average classification accuracy of more than 96%. Future directions of this research include adding more seizure types and improving the classification performance in a broader range of seizure types.

REFERENCES

1. N.I. of Neurological Disorders and S. (US), Seizures and Epilepsy: Hope Through Research. National Institute of Neurological Disorders and Stroke. National Institutes. 2004;4.
2. WHO. Epilepsy. 2019. https://www.who.int/health-topics/epilepsy.
3. Shihata S.S., Abdullah T.S., Alfaidi A.M., Alasmari A.A., Alfaidi T.M., Bifari A.E., Jamal W.H., Rizk H.A. Knowledge, perception and attitudes toward Epilepsy among medical students at King Abdulaziz University, SAGE open medicine. 2021;9:2050312121991248.
4. Paul Y. Various epileptic seizure detection techniques using biomedical signals – a review. Brain informatics. 2018;5(2):1–19.
5. Smith S. EEG in the diagnosis, classification, and management of patients with Epilepsy. Journal of Neurology, Neurosurgery & Psychiatry 2015;76(2):2–7.
6. Lin Lin Lee V., Kar Meng Choo B., Chung Y.S., Kundap U.P., Kumari Y., Shaikh M. et al. Treatment, therapy and management of metabolic Epilepsy – a systematic review. International journal of molecular sciences. 2018;19(3):871.
7. Roy S., Kiral-Kornek I., Harrer S. Deep learning enabled automatic abnormal EEG identification, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE. 2018;2756–2759.
8. Kiral-Kornek I., Roy S., Nurse E., Mashford B., Karoly P., Carroll T., Payne D., Saha S., Baldassano S., O'Brien T. et al. Epileptic seizure prediction using big data and deep learning: toward a mobile system. EBioMedicine. 2018;27:103–111.
9. Golmohammadi M., Ziyabari S., Shah V., de Diego S.L., Obeid I. Picone J. Deep architectures for

automated seizure detection in scalp EEGs, arXiv preprint arXiv. 2017;1712.09776

10. Fisher R.S., Boas W.V.E., Blume W., Elger C., Genton P., Lee P., Engel Jr J. Epileptic seizures and Epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia. 2005;46(4):470–472.
11. Duan L., Lian Z., Chen J., Qiao Y., Miao J., Li M. Classification of epilepsy period based on combination feature extraction methods and spiking swarm intelligent optimization algorithm. Concurrency and Computation: Practice and Experience. 2020;e5550.
12. Guti´errez F.J.M., Roman M.D. Albert D.C. Electro-clinical analysis of epilepsy patients with generalized seizures on adjunctive perampanel treatment. Epilepsy Research. 2020;165:106378.
13. Trinka E., Tsong W., Toupin S., Patten A., Wilson K., Isojarvi J., James D. A systematic review and indirect treatment comparison of perampanel versus brivaracetam as adjunctive therapy in patients with focal-onset seizures with or without secondary generalization, Epilepsy Research. 2020;166:106403.
14. Alarcao S.M., Fonseca M.J. Emotions recognition using EEG signals: A survey. IEEE Transactions on Affective Computing. 2017;10(3):374–393.
15. Houssein E.H., Hassanien A.E., Ismaeel A.A., EEG signals classification for epileptic detection: a review, in: Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing. 2017;1–9.
16. Acharya U.R., Fujita H., Sudarshan V.K., Bhat S., Koh J.E. Application of entropies for automated diagnosis of Epilepsy using EEG signals: A review. Knowledge-based systems. 2015;88:85–96.
17. Harati A., Lopez S., Obeid I., Picone J., Jacobson M., Tobochnik S. The TUH EEG CORPUS: A big data resource for automated EEG interpretation, in: 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE. 2014;1–5.
18. Jaiswal A.K., Banka H. Local pattern transformation based feature extraction techniques for classification of epileptic EEG signals. Biomedical Signal Processing and Control. 2017;34:81–92.
19. Niknazar H., Mousavi S.R., Niknazar M., Mardanlou V., Coelho B.N. Performance analysis of EEG seizure detection features. Epilepsy Research. 2020;167:106483.
20. Yu F. & H.H., Semantic content analysis and annotation of histological images. Computers in Biology and Medicine. 2008;38(6):635–649.
21. Yang S., Li B., Zhang Y., Duan M., Liu S., Zhang Y., Feng X., Tan R., Huang L., Zhou F. Selection of features for patient-independent detection of seizure events using scalp EEG signals. Computers in biology and medicine. 2020;119:103671.
22. Wulandari D.P., Putriz N.G., Suprapto Y.K., Purnami S.W., Juniani A.I. Islamiyah W.R. Epileptic Seizure Detection Based on Bandwidth Features of EEG Signals. Procedia Computer Science. 2019;161:568–576.
23. Deriche M., Arafat S., Al-Insaif S., Siddiqui M. Eigenspace time frequency based features for accurate seizure detection from EEG data. IRBM. 2019;40(2):122–132.
24. Dash D.P., Kolekar M.H., Jha K. Multi-channel EEG based automatic epileptic seizure detection using iterative filtering decomposition and Hidden Markov Model. Computers in biology and medicine. 2020;116:103571.
25. Sharma M., Patel S., Acharya U.R. Automated detection of abnormal EEG signals using localized wavelet filter banks. Pattern Recognition Letters. 2020;133:188–194.
26. Ullah I., Hussain M., Aboalsamh H. et al. An automated system for epilepsy detection using EEG brain signals based on deep learning approach. Expert Systems with Applications. 2018;107:61–71.
27. Nkengfack L.C.D., Tchiotsop D., Atangana R., Louis-Door V., Wolf D. EEG signals analysis for epileptic seizures detection using polynomial transforms, linear discriminant analysis and support vector machines. Biomedical Signal Processing and Control. 2020;62:102141.
28. Alickovic E., Kevric J., Subasi A., Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomedical signal processing and control. 2018;39:94–102.
29. Wei X., Zhou L., Zhang Z., Chen Z., Zhou Y. Early prediction of epileptic seizures using a long-term recurrent convolutional network. Journal of neuroscience methods. 2019;327:108395.
30. Usman S.M., Khalid S., Bashir Z. Epileptic seizure prediction using scalp electroencephalogram signals. Biocybernetics and Biomedical Engineering. 2021;41(1):211–220.
31. Khan H., Marcuse L., Fields M., Swann K., Yener B. Focal onset seizure prediction using convolutional networks. IEEE Transactions on Biomedical Engineering. 2017;65(9):2109–2118.
32. Truong N.D., Nguyen A.D., Kuhlmann L., Bonyadi M.R., Yang J., Ippolito S., Kavehei O. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. Neural Networks. 2018;105:104–111.
33. Obeid I. & Picone J. The temple university hospital EEG data corpus. Frontiers in neuroscience. 2016;10:196.

34. Harati A., Lopez S., Obeid I., Picone J., Jacobson M.P., Tobochnik S. The TUH EEG CORPUS: A big data resource for automated EEG interpretation, in: 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). 2014;1–5. DOI: 10.1109/SPMB.2014.7002953.
35. Shah V., Von Weltin E., Lopez S., McHugh J.R., Veloso L., Golmohammadi M., Obeid I., Picone J. The temple university hospital seizure detection corpus, Frontiers in neuroinformatics. 2018;12:83.
36. Parks T.W. & Burrus C.S. Digital filter design. Wiley-Interscience; 1987.
37. Crochiere R. A general program to perform sampling rate conversion of data by rational ratios Programs for Digital Signal Processing ed AC Schell et al, New York. IEEE Press; 1979.
38. Kubat M., Matwin S. et al., Addressing the curse of imbalanced training sets: one-sided selection, in: Icml. 1997;97:179–186.
39. Yen S.J. & Lee Y.S., Cluster-based undersampling approaches for imbalanced data distributions. Expert Systems with Applications. 2009;36(3):5718–5727.
40. Galar M., Fernandez A., Barrenechea E., Herrera F. EUSBoost: Enhancing ensembles for highly imbalanced datasets by evolutionary undersampling. Pattern recognition. 2013;46(12):3460–3471.
41. Spelmen V.S. & Porkodi R. A review on handling imbalanced data, in: 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT). IEEE. 201;1–11.
42. Vuttipittayamongkol P., Elyan E., Petrovski A., Jayne C. Overlap-based undersampling for improving imbalanced data classification, in: International Conference on Intelligent Data Engineering and Automated Learning. Springer. 2018;689–697.
43. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique, Journal of artificial intelligence research. 2002;16:321–357.
44. Han H., Wang W.Y., Mao B.H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning, in: International conference on intelligent computing, Springer. 2005;878–887.
45. He H., Bai Y., Garcia E.A., Li S., ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence). IEEE. 2008;1322–1328.
46. Rout N., Mishra D., Mallick M.K. Handling imbalanced data: a survey, in: International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications. Springer; 2018;431–443.
47. Breiman L., Random forests, Machine learning. 2001;45(1):5–32.
48. Schindler K., Leung H., Elger C.E., Lehnertz K. Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial EEG. Brain. 2017;130(1):65–77.
49. Kose U. An ant-lion optimizer-trained artificial neural network system for chaotic electroencephalogram (EEG) prediction. Applied Sciences. 2018;8(9):1613.
50. Finnigan S., van Putten M.J. EEG in ischaemic stroke: quantitative EEG can uniquely inform subacute prognoses and clinical management. Clinical neurophysiology. 2013;124(1):10–19.
51. Lodder S.S. & Van Putten M.J. Quantification of the adult EEG background pattern. Clinical neurophysiology. 2013;124(2):228–237.
52. Milton J. & Jung P. Epilepsy as a dynamic disease, Springer Science & Business Media; 2013.
53. Savadkoohi M., Oladunni T., Thompson L. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybernetics and Biomedical Engineering. 2020;40(3):1328–1341.
54. Prathaban B.P., Balasubramanian R. Dynamic learning framework for epileptic seizure prediction using sparsity based EEG Reconstruction with Optimized CNN classifier. Expert Systems with Applications. 2021;170:114533.
55. Wei Z., Zou J., Zhang J., Xu J. Automatic epileptic EEG detection using convolutional neural network with improvements in time-domain. Biomedical Signal Processing and Control. 2019;53:101551.
56. Sudalaimani C., Sivakumaran N., Elizabeth T.T., Rominus V.S. Automated detection of the preseizure state in EEG signal using neural networks, biocybernetics and biomedical engineering. 2019;39(1):160–175.
57. Acharya U.R., Oh S.L., Hagiwara Y., Tan J.H.,Adeli H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, Computers in biology and medicine. 2018;100:270–278.
58. Tsiouris K.M., Pezoulas V.C., Zervakis M., Konitsiotis S., Koutsouris D.D., Fotiadis D.I., A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals, Computers in biology and medicine. 2018;99:24–37.
59. Montans F.J., Chinesta F., Gomez-Bombarelli R., Kutz J.N. Data-driven modeling and learning in science and engineering. Comptes Rendus M´ecanique. 2019;347(11):845–855.
60. Zeiler M.D., Fergus R. Visualizing and understanding convolutional networks, in: European conference on computer vision. Springer; 2014;818–833.
61. Rajaguru H. & Prabhakar S.K., KNN Classifier and K-Means Clustering for Robust Classifica-

tion of Epilepsy from EEG Signals. A Detailed Analysis; 2017.

62. Guler I. & Ubeyli E.D., Multiclass support vector machines for EEG-signals classification, IEEE transactions on information technology in biomedicine. 2007;11(2):117–126.
63. Chen T., Guestrin C., Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016;785–794.
64. Wijnhoven R.G. & de With P. Fast training of object detection using stochastic gradient descent, in: 2010 20th International Conference on Pattern Recognition. IEEE. 2015;424–427.
65. Roy S., Asif U., Tang J., Harrer S. Seizure type classification using EEG signals and machine learning: Setting a benchmark, in: 2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE. 2020;1–6.
66. Lyon D. The Discrete Fourier Transform, part 6: Cross-correlation. J. Object Technol. 2010; 9(2):17–22.
67. Rao K.R., Kim D.N., Hwang J.J. Fast Fourier transform-algorithms and applications, Springer Science & Business Media; 2011.
68. Shaker M.M. EEG waves classifier using wavelet transform and Fourier transform. Brain. 2016;2(3).
69. Zazzaro G., Cuomo S., Martone A., Montaquila R.V., Toraldo G., Pavone L. EEG signal analysis for epileptic seizures detection by applying data mining techniques. Internet of Things. 2019;100048.
70. Boser B.E., Guyon I.M., Vapnik V.N. A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory. 1992;144–152.